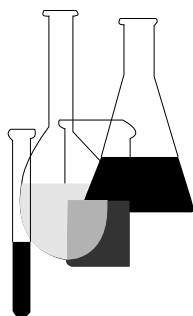




Ecological Effects Test Guidelines

OPPTS 850.2500 Field Testing for Terrestrial Wildlife



“Public Draft”

INTRODUCTION

This guideline is one of a series of test guidelines that have been developed by the Office of Prevention, Pesticides and Toxic Substances, United States Environmental Protection Agency for use in the testing of pesticides and toxic substances, and the development of test data that must be submitted to the Agency for review under Federal regulations.

The Office of Prevention, Pesticides and Toxic Substances (OPPTS) has developed this guideline through a process of harmonization that blended the testing guidance and requirements that existed in the Office of Pollution Prevention and Toxics (OPPT) and appeared in Title 40, Chapter I, Subchapter R of the Code of Federal Regulations (CFR), the Office of Pesticide Programs (OPP) which appeared in publications of the National Technical Information Service (NTIS) and the guidelines published by the Organization for Economic Cooperation and Development (OECD).

The purpose of harmonizing these guidelines into a single set of OPPTS guidelines is to minimize variations among the testing procedures that must be performed to meet the data requirements of the U. S. Environmental Protection Agency under the Toxic Substances Control Act (15 U.S.C. 2601) and the Federal Insecticide, Fungicide and Rodenticide Act (7 U.S.C. 136, *et seq.*).

Public Draft Access Information: This draft guideline is part of a series of related harmonized guidelines that need to be considered as a unit. *For copies:* These guidelines are available electronically from the EPA Public Access Gopher (gopher.epa.gov) under the heading “Environmental Test Methods and Guidelines” or in paper by contacting the OPP Public Docket at (703) 305-5805 or by e-mail: guidelines@epamail.epa.gov.

To Submit Comments: Interested persons are invited to submit comments. By mail: Public Docket and Freedom of Information Section, Office of Pesticide Programs, Field Operations Division (7506C), Environmental Protection Agency, 401 M St. SW., Washington, DC 20460. In person: bring to: Rm. 1132, Crystal Mall #2, 1921 Jefferson Davis Highway, Arlington, VA. Comments may also be submitted electronically by sending electronic mail (e-mail) to: guidelines@epamail.epa.gov.

Final Guideline Release: This guideline is available from the U.S. Government Printing Office, Washington, DC 20402 on *The Federal Bulletin Board*. By modem dial 202-512-1387, telnet and ftp: fedbbs.access.gpo.gov (IP 162.140.64.19), or call 202-512-0135 for disks or paper copies. This guideline is also available electronically in ASCII and PDF (portable document format) from the EPA Public Access Gopher (gopher.epa.gov) under the heading “Environmental Test Methods and Guidelines.”

OPPTS 850.2500 Field testing for terrestrial wildlife.

(a) **Scope**—(1) **Applicability.** This guideline is intended to meet testing requirements of both the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) (7 U.S.C. 136, *et seq.*).

(2) **Background.** The source material used in developing this harmonized OPPTS test guideline is OPP 71–5 Simulated and Actual Field Testing for Mammals and Birds (Pesticide Assessment Guidelines, Subdivision E—Hazard Evaluation; Wildlife and Aquatic Organisms) EPA report 540/09-82-024, 1982. This guideline may be used with OPPTS 850.7100.

(b) **Test standards**—(1) **Test substance.** Unless specified otherwise, data is to be derived from testing conducted with an end-use product. An end-use product may be the applicant's own product or a typical end-use product.

(2) **Test conditions.** The test conditions for conducting a field test should resemble the conditions likely to be encountered under actual use of the product. Specifically, the pesticide should be applied to the site at the rate, frequency, and method specified on the label.

(3) **Endangered species.** Studies should not be conducted in critical habitats or areas containing, or suspected to contain, endangered or threatened plants or animals which may be threatened by the tests to be conducted.

(4) **Residue levels.** When the test substance is applied under field condition testing, and residues of the test substance can be detected in warm-blooded animals, residues should be determined in selected tissues of test organisms and in vegetation, soil, water, sediments, and other appropriate environmental components. If residues of the test substance cannot be detected in warm-blooded animals, the applicant should consult with the Agency before beginning the test.

(5) **Other standards.** Additional standards for conducting field tests are not delineated because of the wide variety of mechanisms by which a pesticide may enter the environment, and because of the great variety of food sources and habitats that may be affected. Any additional standards for conducting these tests will be provided by the Agency in writing following consultation between the applicant or registrant and the Agency, and will take into account the variety of mechanisms, food sources, and habitats mentioned above.

(c) **General**—(1) **Introduction.** (i) Data from full scale terrestrial field studies are required by 40 CFR 158.145 on a case-by-case basis to support the registration of an end-use product intended for outdoor application. Because these studies are complex and costly, the Agency requires

these tests to evaluate only those products that appear to pose significant risks to nontarget wildlife.

(ii) Laboratory tests generally are amenable to a high degree of standardization. In contrast, field study protocols must retain a high degree of flexibility. Because standardization of field studies is made difficult by variables such as chemical mode of action, use pattern, crop type, method of application and species density and diversity, this series of guidelines provides only a general outline for field studies. Specific protocols must be developed as needed and submitted to the Agency for review. However, despite the variability among field studies, several key elements common to most field studies can be identified. This guideline was prepared to identify and discuss these elements as they pertain to terrestrial vertebrates, and to provide a better understanding of the purpose of field studies.

(2) **Objective of field studies.** (i) The purpose of the field study is either to refute the assumption that risks to wildlife will occur under conditions of actual use of the pesticide or to provide some quantification of the risk that *may* occur. The purpose is twofold because the FIFRA requirement to determine unreasonable adverse effects implies the need for a risk-benefit analysis. Thus, if the assumption of risk cannot be refuted, and in order to refine the risk-benefit analysis, field studies should quantify the adverse effects that would occur from actual use of the pesticide.

(ii) A study designed to refute hazard is unusual in biological research. Typically, an investigator is more concerned about concluding with a high degree of confidence that an effect occurred, not that it failed to occur. FIFRA specifies that a pesticide is to be registered only if EPA determines it will not cause unreasonable adverse effects. While the difference between an objective of “will cause” and “will not cause” may seem trivial, it substantially influences study design and the evaluation of data.

(iii) The adverse effects to wildlife that can result from the use of pesticides can be classified as those that affect populations of wildlife and those that affect individuals but not the entire population. Either of these effects may warrant regulatory action, including cancellation or suspension. An adverse effect that results in a reduction in local, regional, or national populations of wildlife species is clearly of great concern. A pesticide that can repeatedly or frequently kill wildlife is also of concern even if these repeated kills may or may not affect long-term populations. The terrestrial field study, accordingly, must be designed to assess both of these types of effects.

(iv) The field study must be designed to provide data that show whether wildlife species will not be affected significantly by a pesticide under normal pesticide use practices. To achieve this objective fully at the population level, detailed knowledge of the population dynamics and

varying environmental conditions for each species potentially at risk must be available. The theoretical aspects of population dynamics are well documented in the literature. However, empirical data are available for only a few species (see paragraph (h)(8) of this guideline). A study designed to provide the needed data would include information on age structure, age-specific survival and reproductive rates, and the nature and form of intrinsic and extrinsic regulatory mechanisms. Such a study, when coupled with the influence of pesticide application on these parameters, would require several, if not many, years in order to begin to give meaningful results. The cost of obtaining such data could make these studies impractical, if not impossible.

(v) The essential question is: How can these studies be performed in a practical, economical manner and still provide data that can show that the pesticide under study will not reduce or limit wildlife populations or repeatedly kill wildlife?

(vi) The question can be answered by examining the potential influence pesticides can have on wildlife. These effects include:

(A) Direct poisoning and death by ingestion, dermal exposure, and/or inhalation.

(B) Sublethal toxic effects indirectly causing death by reducing resistance to other environmental stresses such as diseases, weather, or predators.

(C) Altered behavior such as abandonment of nests or young, change in parental care, or reduction in food consumption.

(D) Reduced food resources or alteration of habitat.

(E) Lowered productivity through fewer eggs laid, reduced litter size, or reduced fertility.

(vii) These effects can manifest themselves in a population through reduced survival and/or lower reproductive success. However, if a field study shows that actual use of a pesticide does not affect survival and/or reproductive success or that only minor changes occur, it would seem reasonable to conclude that the use of the chemical will not significantly impact wildlife. Further, if a field study provides estimates on the magnitude of survival and reproductive effects, it is possible to make reasonable projections on the meaning of the effects to nontarget populations by using available information on the species of concern and basic theories of population dynamics. While less than ideal, field studies that collect information on survival and reproductive effects and use these data to address population parameters should provide a reasonable basis for evaluating potential impacts. This is not to imply that effects on populations are

the only concern, however, a study adequate to assess these effect will also assess the degree of risk to individual wildlife.

(viii) This guideline emphasizes avian and mammalian wildlife. The Agency is also concerned about other terrestrial organisms such as nontarget plants, invertebrates, amphibians, and reptiles. Plants and invertebrates are excluded here from direct study, except as sources of food or pesticides to wildlife. Testing guidelines for nontarget insects and plants are in Groups C and D, respectively, of the 850 guideline series. Established protocols, especially for acute and chronic toxicity testing, are available for birds and mammals, but not for reptiles and amphibians. Further, the Agency assumes that “protection” for reptiles and amphibians is provided through the risk assessment process for birds and mammals. Occasionally, however, it may be necessary to adapt these field techniques to apply specifically to reptiles and/or amphibians.

(3) **General approach.** (i) Field studies required to support registration have evolved into two types, screening and definitive. The types of studies required depends on the available data on the chemical in question. If the available information is limited to laboratory toxicity data on a limited number of species, coupled with effective environmental concentrations (EECs), a screening field study may be appropriate to determine if impacts are occurring and, if so, to what species. If a screening study indicates impacts are occurring, or if other available data suggest that deleterious effects have occurred or are extremely likely, the study design should be quantitative, evaluating the magnitude of the impacts in a definitive study. For some chemicals it may be appropriate to proceed directly to a definitive study without the screening phase. Careful consideration needs to be given to the likelihood of impacts occurring in order to determine which approach to use. When there is insufficient information to indicate which species are at risk in the field but available data strongly suggest adverse effects will occur, it may be appropriate for a field study to begin with the general approach of a screening study, followed by a quantitative phase that focuses on the species affected in the screening phase. In certain instances there may be sufficient data and information for the Agency to decide additional testing including field testing is not necessary prior to a regulatory action.

(ii) Design differs between the screening study and the definitive study. If the objective of the study is to determine if impacts are occurring, “pass-fail” methods can evaluate whether or not animals are being stressed by the application. These methods may include carcass searching, residue analysis of species collected on study plots, residue analyses of wildlife food sources found in and adjacent to the area of application, behavioral observations, and enzyme analysis. At the quantitative level (definitive study), the objectives should include estimating the magnitude of acute or secondary mortality caused by the application, the existence and extent of reproductive effects, and the influence of pesticide use on the

survival of species of concern. Methods that can be used to address these objectives include mark-recapture, radio telemetry, line transect sampling, nest monitoring, territory mapping, and measuring young to adult ratios.

(4) **Sampling and experimental design.** While examples of acceptable experimental designs are given, it is beyond the purpose of this guideline to cover the fundamentals of this topic. Paragraph (f) of this guideline provides a general outline for a field study protocol to be submitted to the Agency for review. Paragraphs (d) and (e) of this guideline outline points to be considered in designing screening and definitive field studies. As stated under paragraph (c)(1) of this guideline, specific protocols must be developed on a case-by-case basis and submitted to the Agency for review.

(d) **Screening study—(1) Objective and scope.** (i) The screening study is designed primarily to demonstrate that the hazard suggested by lower tier laboratory or pen studies does not exist under actual use conditions. The interpretations of screening study results, in most cases, are limited to “effect” versus “no-effect” determinations. If the study indicates that the pesticide has caused little or no detectable adverse effect, it may be reasonable to conclude that potential adverse effects are minor. When effects are demonstrated, it may be necessary to determine the magnitude of the effects, thus requiring additional testing if pesticide registration or continued registration is still desired. Therefore, when information already available shows that a product has caused adverse effects under normal use conditions, the screening study may be of limited value. It may be appropriate to proceed directly to a definitive field study in cases where analysis of laboratory or other data strongly suggest that adverse effects are likely to occur, and are unlikely to be attenuated by field use conditions.

(ii) Screening studies are limited to addressing the potential for acute toxic effects, such as direct poisoning and death, and sublethal toxic effects potentially affecting behavior and/or survival. In most instances, a screening study would not address chronic effects, such as reduced reproduction, or effects such as changes in density or diversity of populations.

(iii) Further laboratory and/or pen studies may be useful prior to proceeding to the field, or may be necessary to interpret results of the field study. For example, additional toxicity data on species that are expected to be exposed from the proposed use pattern may indicate which species are more susceptible to the pesticide, allowing the study to be designed to monitor those species in greater depth as well as to provide insight into field results that show some species were affected more than others, and additional laboratory studies may be unavoidable. If residue concentrations in resident species are being used to indicate potential problems, the relationship between tissue levels and the doses that causes adverse effects must be estimated. If secondary poisoning is of concern, feeding secondary

consumers (held in captivity) prey items collected in the field following the application can be useful to evaluate this potential exposure route. Laboratory toxicity tests for secondary consumers coupled with residue analysis of prey items can indicate the potential for secondary poisoning of nontarget species. In designing field studies, the utility of laboratory and/or pen tests should not be neglected, and where appropriate their use is encouraged.

(2) **Geographic area selection.** (i) The selection of geographical areas for evaluating pesticide impacts on wildlife can be difficult, particularly for pesticides to be used on crops grown over large and diverse areas. Studies should be performed in each biogeographic area where the pesticide could be used. While this approach may be practical for uses restricted to localized areas and conditions, many uses (e.g., corn, soybeans, alfalfa) would require an inordinate number of studies in different geographic areas, due to the diversity and variability in wildlife species and habitats involved. To keep the number of geographic areas at a manageable level while still accomplishing the purpose of the field study, geographical area selection should be biased toward situations likely to present the greatest risk. If hazards appear to be low under these conditions, it can be reasonably concluded that impacts under less severe conditions would be minor.

(ii) A careful review of the species and habitats in the various geographical areas where the pesticide could be used is necessary to identify the areas of highest concern. A sound understanding of the biology of the species that are found in association with the potential use sites is essential. Identifying these areas may require an extensive literature review and consultation with experts familiar with the areas and species of concern. The study area selected should be frequented by those species that would have high exposure, based on their feeding or other behavioral aspects. If exposure and fate (e.g., degradation) parameters vary geographically, study area selection also should be biased towards maximizing residues available to wildlife. In some circumstances preliminary monitoring of candidate areas may be necessary to determine which should be selected for detailed study,

(3) **Study site selection.** (i) Selection of study sites within each geographic area also is extremely important in designing field studies. Study sites should be randomly selected throughout the study area. While this approach may be practical for some areas such as rangeland or large contiguous crops, random selection would require a large number of sites to provide a representative sample due to the diversity and variability in wildlife species and habitats in most areas. The cost and time requirements of such studies would be unreasonable. To maximize the hazard, the sites selected should have associated species that would be at highest risk from the application, as well as a good diversity of species to serve as indicators for other species not present at that specific location. The choice of study

sites that are as similar as possible in terms of abundance, diversity, and associated habitat will facilitate an analysis of the results.

(ii) Under some circumstances, it may be difficult to decide beforehand which species are likely to be at highest risk. In most cases, field surveys of a number of sites may be needed to identify which sites should be selected for detailed study. Even when high risk species can be identified, preliminary surveys may be needed to determine which sites have adequate numbers of the high risk species as well as a good diversity of other species.

(iii) In general, study sites should be selected from what is considered to be a typical application area, but at the same time, study sites should contain the widest possible diversity and density of wildlife species. Identifying potential study sites may require consultation with experts familiar with the areas where studies are proposed, and preliminary sampling.

(iv) In the initial evaluation of potential study sites, edge effect may indicate which sites support the larger and more varied wildlife populations. As stated by Aldo Leopold (see paragraph (h)(23) of this guideline), “The potential density of game of low radius requiring two or more types is, within ordinary limits, proportional to the sum of the type peripheries.” (Type is defined as the various segments of an animal’s environment used for food, cover, or other requirements.) If study sites are selected to maximize edge effect the potential for high density and diversity should be increased. One quantitative measure of edge and edge effect (see paragraph (h)(15) of this guideline) is the distances around individual plant communities in relation to the unit area of the community. Population densities, in general, are positively related to the number of feet of edge per unit area of community. Study sites chosen to maximize the ratio of edge to core may serve to indicate sites with higher densities and diversities of wildlife species.

(v) While this ratio can be helpful in selecting study sites, the other characteristics of edge should not be neglected in screening potential study sites. Density and diversity of wildlife species are also influenced by the variety in the composition and arrangement of the edge component cover types and its width. Also, interspersions (the plant types and their association with one another) influences densities of wildlife species. The edge effect is the sum of all the characteristics of edge and hence each component needs to be considered. An agricultural field with a relative high edge to core ratio may not have as high a density and diversity as one with a lower ratio but greater variety, width, and interspersions. In general, edge characteristics can be used to screen potential study sites; however, preliminary sampling of prospective study sites will be needed to identify study sites with adequate density and diversity of wildlife species.

(4) **Number of sites.** (i) The number of sites needed can be estimated using the binomial theorem. Briefly, the rationale is that for each study site there are two possible outcomes, either effect or no-effect. Trials of this type are known as binomial trials and, when repeated, the results will approximate a binomial distribution. In this case, to use the binomial theorem, it is necessary first to define the expected probabilities that birds or mammals on a site are affected or not affected, after which the probability of the discrete binomial random variable x for n replications can be used to determine the minimum number of sites at a certain level of significance.

(ii) For purposes of illustration, a problem exists if some specific mortality rate or level of some other variable occurs on more than 20 percent of the potential application sites. Translated into binomial probabilities, there is a 0.2 probability of a site showing an effect and a 0.8 probability of a site not showing an effect. Therefore, if the results from the field trial show that the number of sites affected is significantly lower than 0.2 n , it can be concluded that potential impacts will be below the stated level of concern.

(iii) To calculate the minimum number of sites necessary to show a significant difference between the observed and expected, the following formula for the probability of the binomial random variable x can be used (see paragraph (h)(36) of this guideline):

$$P(x) = \binom{n}{x} p^x q^{n-x}$$

where x = number of sites showing effects, n = number of sites, p = probability of a site showing an effect, and q = probability of a site not showing an effect.

Then, solving for n , when $x = 0$, i.e.,

$$P(x = 0) = \binom{n}{0} q^n$$

Let $P(x = 0) = \alpha$, then

$$\alpha = q^n$$

$$\log \alpha = n \log q$$

$$n = \log \alpha / \log q$$

(iv) The minimum number of sites can be determined using this formula. Continuing with the discussion example of 20 percent occurrence of an effect as a level of concern (i.e., a 0.2 probability of an affected

site, a 0.8 probability of a nonaffected site, and a 0.05 level of significance), n would be:

$$n = \log 0.05 / \log 0.8$$

$$n = 13.43$$

Therefore, 14 is the minimum number of sites needed such that the probability is not greater than 0.05 that all sites surveyed would be unaffected. In other words, if 20 percent of the application sites are actually affected, there is only a 5 percent chance of finding all 14 sites unaffected when $n = 14$. Moreover, if 20 percent of the application sites are actually affected, we expect to find 1, 2, 3, and 4 sites affected with probabilities of 0.15, 0.25, 0.25, and 0.17, respectively, when $n = 14$.

(v) Under many circumstances, conducting this number of replications may not be practical. However, the number of sites can be reduced if site selection is biased toward hazard. While arguable, it seems logical that if the worst cases are sampled, a less stringent level of significance could be accepted. While this must be determined on a case-by-case basis, the Agency believes a minimum acceptable level of significance under worst-case conditions is 0.2 rather than 0.05 under average or normal use conditions. At this level, eight sites showing no effect would be required to conclude at the 0.2 level of significance that the effect occurred on less than 20 percent of the application sites or—there is less than a 20 percent chance that all eight sites will be judged unaffected when $n = 8$ sites. Under some circumstances, this may not seem adequately protective. It should be noted, however, that based on this same design, it could be concluded that, at the 0.1 level of significance, effects occur on less than 30 percent of the application sites, and at the 0.05 level of significance, effects occur on less than 40 percent of the application sites. Hence, with eight sites, it could be concluded with a relatively high degree of confidence that effects would occur on less than 40 percent of the application sites. Because worst-case study sites were used, the Agency could have additional confidence that adverse effects would occur on less than 20 percent of all normal application sites.

(vi) Under some circumstances, particularly if endangered species could be exposed from the proposed use, additional replication may be desirable. Under these conditions, a high degree of confidence that an effect was a rare occurrence would be required. (Under no circumstances should field studies on pesticides be conducted in areas where endangered species could be exposed.)

(vii) The calculations under paragraph (d)(4)(iii) of this guideline are for when $x = 0$, no effects are observed on any site. A similar approach can be used to estimate the number of sites necessary to show a significant result for a critical value of $x > 0$. Again the formula for the probability of the binomial random variable can be used summing the probabilities of x and all outcomes $< x$. Then by using increasing values of n , the num-

ber of replications required to show statistical significance may be determined for a given level of significance for individual x values. That is:

$$P(X \leq r) = \sum_{x=0}^r \binom{n}{x} p^x q^{n-x}$$

The minimum value of n occurs when

$$P(X \leq r) = \alpha \text{ level}$$

Continuing the previous example, the following Table 1. gives the results for $x \leq 1$ and $x \leq 2$ for the previously defined acceptable occurrence level of effect (i.e., a 0.2 probability of an affected site, a 0.8 probability of nonaffected site). From the table, the minimum number of sites needed when the critical value for x is set at 1, to conclude that (at a 0.2 level of significance) effects are occurring below levels of concern is 14. If $x \leq 2$, 21 are needed in order to have an equivalent criterion. As can be seen, as x (the number of sites with effects) increases, the number of sites required to show a statistical significance becomes inordinately large.

Table 1.—Probabilities for Binomial Random Variable with $p = 0.2$ for $x \leq 1$ and ≤ 2 as a Function of the Number of Sites (N).

N	$P(x \leq 1)$	$P(x \leq 2)$
8	0.5033	0.7969
9	0.4362	0.7382
10	0.3758	0.6778
11	0.3321	0.6174
12	0.2749	0.5583
13	0.2336	0.5017
14	0.1979	0.4481
15	0.1671	0.3980
16	0.1407	0.3518
17	0.1182	0.3096
18	0.0991	0.2713
19	0.0827	0.2369
20	0.0692	0.2061
21	0.0576	0.1787

(viii) When the probability of an affected site is 0.2, application of the rule of zero-observed-affected-sites results in a declaration of no-effect 16.8 and 13.4 percent of the time for samples of size 8 and 9, respectively. It also results in a declaration of no-effect 43.1 and 38.7 percent of the time for samples of size 8 and 9, respectively, when the probability of an affected site is 0.1, a value less than the criterion probability. Application of the rule of zero-observed-affected-sites for a declaration of no-effect means that a study is considered to be negative (shows no effect) if, and only if, none of the sites show effects.

(ix) Under any condition, it is extremely important with the binomial approach to define the critical or threshold level for an effect, and to be sure that the methods used are sensitive enough to detect an effect should one occur. These assessments depend upon the species potentially at risk as well as the parameter being sampled. It should be noted that the measure of effect is not limited to mortality. Other parameters, such as residue or enzyme levels, can be used. Whatever parameters are used, defining the criteria level for an effect is extremely important, and when designing studies this issue should be considered carefully.

(x) Using this approach, control (reference) sites are not an absolute necessity. While the Agency encourages their use, in some cases the additional information gained from the control sites for a screening study may not justify the additional effort required. In most instances, control sites would serve to protect from erroneously attributing effects due to other causes to the pesticide. However, for most chemicals, this can be avoided by employing methods, such as residue analysis and/or cholinesterase (ChE) inhibition tests, that can be used to indicate if the pesticide contributed to the observed effect. Further, because studies have shown that it is a relatively rare event to locate dead or sick animals in the wild except under unusual conditions (see paragraph (h)(18) of this guideline), it is unlikely to find dead animals that were killed by something other than the pesticide being tested.

(xi) Nevertheless, controls may be necessary when reliable methods to confirm the cause of effect are not available. In these cases the binomial design can be modified to a paired-plot binomial design, with a treatment plot and a comparable control plot for each study site within an area. When critical levels of effect and occurrence are defined, the binomial theorem can be used for sample size determination, which gives 8 site pairs (16 paired plots) showing less than a defined difference between plots to conclude at the 0.2 level of significance that the effect occurred on less than 20 percent of the application sites. Alternatively, a quantitative difference or, preferably, ratio of treated to control responses could be used to test for a treatment effect on each of the measured response variables. (This is discussed further under paragraph (e)(2) of this guideline.)

(5) **Size of study sites.** Study sites must be large enough to provide adequate samples. The size is dependent on the methods used, the sensitivity required, and the density and diversity of species and their ranges. In some cases, particularly with slow-acting poisons or where species at high risk have relatively large home ranges, areas several times larger than the treatment area may need to be examined. In some circumstances, several fields in an area may be included in a single study site to account for wide-ranging species of lower densities. Except in the unusual circumstance where fields are extremely large (e.g., forested and range areas), the study site should never be less than an individual field and the surrounding area. The nature of the surrounding area is discussed further

under individual methods. Another consideration is the distance between study sites—in general, sites should be separated adequately to ensure independence, which is dependent mainly on the range of the species that could be exposed.

(6) **Chemical application.** (i) Consideration should be given to application rates and methods. In general, the test conditions should resemble the conditions likely to be encountered under actual use of the product. In most instances the pesticide should be applied at maximum use rates and frequencies specified on the label.

(ii) If more than one application method is specified on the label, the method that maximizes exposure of nontarget species should be used. This evaluation should relate wildlife utilization of the area to exposure. For example, if the crop is one that is used by avian species as preferred nesting areas, feeding areas, or cover, ground application may be the method that maximizes exposure. However, if it is a crop with low utilization by wildlife species, but with high utilization of its edges, aerial application where drift could increase exposure may be more appropriate. In any case, the method of application used must be consistent with the label.

(iii) Equipment used may influence potential exposure of nontarget species. There is a diversity of types of farming equipment that, depending on the particular use pattern involved, could influence exposure. For example, for pesticides applied in-furrow at planting there are several types of covering devices employed on seeders, such as drag chains, drag bars, scraper blades, steel presswheels, etc., in which the efficiency may vary for covering the pesticide. In general, the various equipment normally used for the particular pesticide application has to be evaluated to estimate the potential influence of equipment choice on exposure. In some instances, preliminary tests may be required to estimate which method and equipment poses the highest exposure.

(7) **Methods.** (i) This section provides a general outline of methods appropriate for use in a screening field study and indicates some of their limitations. While methods described have been found to be most useful, a screening study is not limited to these methods. If other methods are more appropriate, their use is encouraged. Because procedures should be adapted to specific situations, the outlines presented should not be interpreted as strict protocols. Normally, different methods will be combined to evaluate potential impacts. Due to the indefinite number of variables and the unpredictability of wild animals, even normally reliable procedures can sometimes prove inadequate.

(ii) The methods used in a screening study address exposure by monitoring overt signs of toxicity such as mortality or behavioral modifications, or by evaluating parameters that indicate animals are under stress, such as residue concentrations in tissues or degree of enzyme inhibition.

Measurements of density and diversity of species are needed to aid in evaluating the results. The following methods can be useful for screening studies.

(A) **Carcass searches.** (1) Searching for dead or moribund wildlife is a basic method used in field studies to evaluate the impact of pesticides on nontarget species. Carcass searches can roughly indicate the magnitude of kills when adequate areas are searched and the reliability of the search is documented. This latter point is extremely important. Rosene and Lay (under paragraph (h)(30) of this guideline) indicated that finding even a few dead animals suggests that there has been considerable mortality; failure to find carcasses is poor evidence that no mortality has occurred. The reliability of the search is based upon the percentage of carcasses recovered by searchers and the rate of disappearance. By knowing the reliability, the significance of failure to find carcasses can be assessed and the extent of the kill estimated.

(2) Finding dead animals is seldom easy, even if every animal on a site is killed. For example, three breeding pairs of small birds per acre is considered a large population (see paragraph (h)(18) of this guideline), and under average cover conditions, a small bird is difficult to detect. Small mammals may be more abundant but, due to their typically secretive habits, they are more likely to die under cover and can be even more difficult to find than birds. Carcass searching specifically for mammals should be attempted only when cover conditions permit reasonable search efficiency. However, any vertebrate carcasses found should be collected, even if the search is oriented primarily to one taxon.

(3) Because results may be biased by scavenging and failure to find carcasses, the sensitivity of this procedure should be determined. Under conditions of heavy cover and/or high scavenger removal, other methods may be more appropriate.

(4) There are no standard procedures for carcass searches. Paragraph (g) of this guideline outlines practices that have been used typically and should be considered in designing searches.

(B) **Radio telemetry.** (1) Radio telemetry has been found to be extremely useful for monitoring mortality and other impacts caused by pesticide exposure of wildlife. Advances in miniaturizing electronic equipment have made it feasible to track most vertebrate animals. Transmitters weighing a few grams have been used to track species as small as mice. Cochran's excellent summary of this technique provides additional details (see paragraph (h)(5) of this guideline).

(2) Radio telemetry has the advantages of providing information on the fate of individual animals following a pesticide application and of facilitating carcass recovery for determining the cause of death. Although the initial cost of this technique may be more than for other methods,

the increase in information obtained under some circumstances can more than justify the cost. The method is particularly useful with less common or wide-ranging species.

(3) In addition to mortality, radio telemetry can be used to monitor behavioral modification as well as physiological changes. Automatic radio-tracking systems permit continual surveillance of the location of animals (see paragraph (h)(5) of this guideline), which can be used to provide insight into behavioral changes such as nest abandonment, desertion of young, or decreases in activities such as flying or feeding. Radio telemetry equipment is also available for the transmission of physiological data such as heart rates or breathing rates (see paragraph (h)(25) of this guideline).

(4) While this technique can provide very useful information on impacts of pesticides to wildlife, other points need to be considered in addition to cost. Capturing animals alive and unharmed requires time, skill, and motivation. For the method to be consistently successful, the investigator must be thoroughly familiar with the habits of the species under study and with the various capture methods that can be used. Even for the most experienced investigator, adequate sample sizes can be difficult to obtain.

(5) Adequate sample size is very important. The binomial theorem can be used to estimate minimum sample size per site, if the question is limited to mortality. Briefly stated, to be sure that nontarget species are not being affected by environmental concentrations greater than, for example, an LC20, the expected binomial probabilities would be 0.2 for mortality and 0.8 for nonmortality. Depending on the level of significance, 8 ($\alpha = 0.2$) to 14 ($\alpha = 0.05$) individuals would need to be monitored per site (see paragraph (d)(4) of this guideline for further details on these calculations). However, since the LC20 may differ between species, 8 to 14 individuals would be required for each species, unless laboratory tests have documented relative species sensitivity. Further complications can arise if radiotagged animals leave the area or if the movements of individuals limit their exposure. If these complications occur at relatively low rates, a few additional radio-tagged animals may be sufficient to overcome these problems.

(C) Tests of ChE inhibition. (1) Measuring ChE concentrations in animal tissues has been found to be a very useful field technique for evaluating exposure of nontarget animals to ChE-inhibiting-chemicals (see paragraphs (h)(18) and (h)(19) of this guideline). These chemicals, including organophosphates and carbamates, affect the synaptic transmission in the cholinergic parts of the nervous system by binding to the active site of acetylcholinesterase (AChE), which normally hydrolyzes the neurotransmitter acetylcholine. Thus, ChE inhibitors permit excessive acetylcholine accumulation at synapses, thereby inhibiting the normal cessation of nerve impulses (see paragraphs (h)(6) and (h)(27) of this guideline).

(2) The depression of AChE activity, when measured and compared to controls, can indicate the degree to which an animal is affected. Brain ChE depression of >50 percent in birds has been found sufficient to assume that death is pesticide-related (see paragraph (h)(24) of this guideline). Depressions of more than 70 percent are often found in dead birds poisoned by these chemicals (see paragraphs (h)(1), (h)(2)), and (h)(33) of this guideline), although some individual birds with less than 50 percent inhibition may die. A 20 percent depression of brain ChE has been suggested as an indication of exposure (see paragraph (h)(24) of this guideline). ChE concentrations in blood can also be used to indicate exposure, avoiding the necessity of sacrificing the animal. However, blood ChE concentrations are influenced more by environmental and physiological factors than are brain ChE concentrations. Because ChE activity varies among species, the degree of depression must be based on an estimated normal value for concurrently tested controls of the species potentially at risk. Because of this difference between species, each case must be considered unique (see paragraph (h)(19) of this guideline).

(3) Although there are several calorimetric methods for determining ChE activity, the general methods are similar. Brain tissues (or blood samples) are taken and analyzed for ChE concentrations. Comparisons are made between pre- and posttreatment and between treated and untreated areas. It is important to ensure that untreated controls have not been exposed to any ChE inhibitors. It also should be noted that absolute enzyme levels in the literature are derived from various different, although similar, methods and are reported in different ways. For example, Ludke et al. (see paragraph (h)(24) of this guideline) used a modification of the method reported in paragraph (h)(13) of this guideline and reported results of ChE activity as nanomoles of acetylthiocholine iodide hydrolyzed per minute per milligram of protein, whereas Bunyan et al. (see paragraph (h)(1) of this guideline) used their own colorimetric method (in addition to a pH change method) and reported the results as micromoles of acetylcholine hydrolyzed per hour per milligram of protein. Therefore, without a tightly standardized method, it is necessary to use concurrent controls of the same species obtained from the general vicinity (but untreated) of the exposed birds, rather than literature values. Because of the greater variation in plasma ChE levels than for brain, more controls are necessary to evaluate blood samples.

(4) Tests for ChE activity can be used to help confirm cause of death and monitor levels of exposure. In the latter case, 5 to 10 individuals of each species are collected before treatment and at periodic intervals following treatment. Mean inhibition of 20 percent or more is considered an indication of exposure to a ChE inhibitor. Confirmation of cause of death may be determined by analyzing brain tissue from wildlife found dead following treatment and comparing the activity with controls. Inhibition of 50 percent or more is considered strongly presumptive evidence that

mortality was caused by a ChE-inhibiting compound. The cause-effect relationship can be further supported by chemical analysis of the contents of the digestive tract or other tissues for the chemical in question.

(5) For this technique to provide accurate information, prompt collection and proper preservation of specimens are essential. ChE concentrations in tissues are influenced by time since death, ambient temperatures, and whether or not reversible ChE inhibitors are being investigated. Therefore, the response of postmortem brain ChE to ambient conditions can seriously affect diagnosis of antiChE poisoning. Samples must be collected shortly after death and frozen immediately to halt changes in tissue or enzyme-inhibitor complexes. A technique for field monitoring and diagnosis of acute poisoning of avian species has been reviewed, discussing sample collection, sample numbers, preservation procedures, and sources of error (see paragraph (h)(19) of this guideline).

(D) **Residue analysis.** (1) Residue analyses of wildlife food sources provide information about the level and duration of pesticide exposure. Residue analysis of animal tissues also can indicate actual exposure levels. If the relationship between tissue concentrations and toxic effects is known for the species in question, residue analyses can provide a measure of the degree to which animals are affected. For this application of residue analyses, laboratory trials are necessary to establish the relationship between residue levels and toxicity. In addition to death, these laboratory trials should include such signs as anorexia, asthenia, asynergy, or ataxia. For chemicals that are readily metabolized by vertebrates, residue analysis may not be appropriate for diagnostic purposes. With many pesticides, it will be necessary to analyze for residues of active metabolites also.

(2) For determining residues on wildlife food sources, the investigator should collect samples of insects, seeds, leafy parts of plants, etc., immediately after pesticide application and at periods thereafter. Samples should be analyzed for the chemical to determine potential exposure rate and duration. The application method needs to be considered in determining where to take samples. If drift is likely, samples should be taken from habitats surrounding the treatment sites as well as in the treated fields. Because analysis can be costly, the investigator should consider carefully the number of samples necessary to provide adequate data. Where feasible, samples from different locations within a site should not be pooled. Separate analysis of samples can provide data on the range and variability of exposure as well as mean levels.

(3) When residue analysis is used to evaluate exposure in nontarget animals, the tissues selected for analysis differ depending on the purpose. Heinz et al. (see paragraph (h)(18) of this guideline) indicated that for many chemicals, residues in brains of birds and mammals can be used to determine if death is pesticide-related. The authors believe that sublethal exposure is judged better from residues in other tissues. Therefore, they

proposed that analyses of whole body homogenates should be used to quantify the body burden of a pesticide. If this is not feasible, analysis of muscle tissue is suggested, because muscle residues reflect body burden more nearly than those of any other tissue, and the amount of muscle tissue is not unduly large. For persistent chemicals, it has been suggested that residues in liver and fat tissues could be misleading for determining acute body burdens. Liver is a processing organ and its residue level largely represents current availability of the chemical. Residues in fat are greatly affected by changes in the amount of body fat, and are not dependable indicators of body burden of the chemical; however, for some chemicals, liver, fat, or other tissues may be good qualitative indicators that exposure did occur. In general, laboratory trials or data gathered in metabolism or other studies may be necessary to determine which tissues can provide the most useful information. Residue analysis of eggs taken from nests in treatment areas can indicate the degree of contamination that a treatment has caused, as well as possible reproductive effects of the treatment.

(4) Two approaches may be used to determine the number of samples to be collected. Frequently, residue samples will be collected to establish a mean value and confidence limits. To determine the number of samples to collect, it is necessary to estimate the standard deviation and to set an arbitrary limit from the mean value that is acceptable. Although the mean value does not need to be estimated, it is also necessary to have some idea of the mean so that the standard deviation can be estimated and the limit can be set. The formula for the number of samples, presented under paragraph (h)(34) of this guideline, is

$$n = 4\sigma^2/L^2$$

for 95 percent probability, where σ is the standard deviation and L is the allowable limit around the mean. For example, to calculate the residue concentrations on vegetation within ± 10 ppm, with an estimate of the standard deviation of 20 ppm, then $n = 4(20)^2/(10)^2$ means 16 samples are required to have a 95 percent probability that the sample mean value will be within ± 10 ppm of the true mean.

(5) In some situations, there may be little information useful for estimating the standard deviation, or the standard deviation may be rather large, thus requiring a very large sample size. For some types of samples, such as residues in nontarget wildlife carcasses, the sample size cannot be increased to permit more precision. The mean value of a parameter certainly has utility, but it also is very important to establish confidence limits around the mean. In general, the Agency will use the 95 percent confidence limits (usually the upper boundary, as in the case of residues) in the assessment of the data. This approach will substantially reduce the impact of outliers but will still incorporate the range of reasonable values into the assessment. In addition, the use of confidence limits reduces the necessity for taking a large number of samples. Because the width of the

confidence intervals decreases with increasing sample sizes, investigators should take samples that are as large as feasible.

(6) Since the sample size will nearly always be less than 30, the calculation of confidence limits should be based on Student's t-distribution. The t values are derived from tables available in most statistics books, and the 95 percent confidence limits are:

$$\bar{x} \pm (t_{.05})(s \div \sqrt{n})$$

where s is the standard deviation estimated from the sample of size n.

(7) Alternatively, the binomial approach may be used for determining if residues, typically in collection of live nontarget animals, exceed a particular threshold value that indicates an effect. The required sample size is the same as presented for the binomial approach in determining the number of study sites. Specifically, in the preceding example a minimum of 8 samples with none exceeding the threshold value or 14 samples with one or none exceeding the threshold value indicates no effect at $p = 0.2$ in 20 percent of the samples. This approach requires the establishment of threshold values which are determined on a case-by-case basis. In general, residues reflecting an LC20 level of exposure would seem to be a maximum acceptable effect concentration for a screening study. An LC50 should be determined in the laboratory for each species analyzed for residues, after which a group of animals would be exposed to an LC20 concentration to determine the mean threshold concentration of residues. Since this approach is impractical for a screening study, it is suggested that the mean residue concentration in bobwhite and/or mallards exposed to an LC20 dietary concentration would provide an indication of threshold levels.

(8) The number and timing of collection periods must be considered and should be based on the persistence of the specific chemical under study. Where persistence in the field has not been adequately determined, it may be necessary to sample at regular intervals (e.g., days 0, 1, 3, 7, 14, 28, 56) to provide data on degradation rates.

(E) **Behavioral observations.** Observations of behavior sometimes can be an extremely important indicator of treatment effects. Such observations might include characteristic signs of toxicity or behavioral changes seen in test animals exposed to the pesticide in the laboratory. Other abnormal behavior (e.g., territorial males abruptly ceasing singing, birds not feeding, reduced avoidance of humans) also may be important.

(F) **Density and diversity estimates.** (1) It is necessary to know the number of individuals and variety of species on and around a study site in order to indicate which species could have been exposed and to aid in evaluating the significance of mortalities or other findings. In addition,

preliminary information on density and diversity is necessary for site selection and to determine the size of study sites. Under some circumstances, comparisons of density estimates between treatment and control sites, or between before and after treatments, may be used to indicate pesticide impacts. In general, the usefulness of these comparisons is limited in a screening study due to the relatively small acreage involved. If mortality occurs, replacement from outside is likely to be so rapid that losses are replaced before censuses are completed. Seasonal changes, such as migration, molt, or incubation, that can affect real or apparent densities must be considered.

(2) Several techniques may be used to estimate the density and diversity of wildlife species, including counts of animal signs, catch-per-unit effort, mark-recapture, and line-transect sampling. Although the methods selected depend on the species of concern, for the screening field test line transect methods are likely to be the most useful for birds.

(3) The major advantage of line-transect sampling is that it is relatively easy to use in the field once a proper sample of lines has been chosen. However, line-transect-sampling is not applicable to all species, particularly those that are not easily observed. Individuals using line-transects must be extremely competent in species identification.

(4) In the line transect method, an observer walks a distance (L) across an area in nonintersecting and nonoverlapping lines, counting the number of animals sighted and/or heard (N), and recording one or more of the following statistics at the time of first observation:

(i) Radial distance from observer to animal.

(ii) Right-angle distance from the animal sighted or heard to the path of the observer or angle of sighting from the observer's path to the point at which the animal was first sighted or heard.

(5) Although the field procedures are simple, they must be understood adequately and implemented well to obtain good estimates of density (see paragraph (h)(3) of this guideline). The authors provide a thorough review of the theory and design of line-transect-sampling and this monograph should be reviewed for details.

(6) For mammals, density and diversity estimates from capture data may be the most practical for a screening study. There are several ways of estimating the populations from capture data, some relatively simple, that may provide adequate information for a screening study. Davis and Winstead (see paragraph (h)(7) of this guideline), review the various methods available, explaining their advantages and disadvantages.

(8) **Interpretation of results.** (i) The numerous variables involved in field studies makes a meaningful discussion of the interpretation of re-

sults somewhat tenuous, particularly with the almost inexhaustible array of results that could occur. Each study must be considered unique and therefore will require a case-by-case analysis that incorporates not only the actual study but other relevant information that is available.

(ii) In general, the results of the screening field study should provide information on acute poisoning and potential sublethal effects as suggested by enzyme, residue, or other measurements. In addition, information has to be developed on the density and diversity of species on the study sites as well as the sensitivity of the methods used. If no effects are detected, assuming that the methods used were adequate to detect levels of concern and that the species on the study site represent a good cross-section of the nontarget species expected to be at risk, the potential hazard indicated by lower tier tests is refuted. Unless other hazards (e.g., reproduction) are still of concern, additional tests would not normally be necessary. However, if an effect is detected on one or more study sites at rates equal to or greater than concern levels, the hazard has not been refuted and additional tests may be necessary.

(iii) In interpreting if an effect has occurred in the context of the binomial approach, care must be employed not to assume a level of precision in results that does not exist. Most detectable effects will exceed the concern level for some methods used in these studies, due to the inherent variability in the data collected to estimate the level of impact, particularly when minimum sample sizes and areas are used. In some instances in interpreting results it may be appropriate to use confidence limits of data collected (or another measure of dispersion) to evaluate if concern levels are exceeded. For example, when density estimates are used to estimate percent mortality using the number of dead animals found during carcass searching, the upper and lower confidence limits of the density estimate may be more appropriate than the average, particularly when variability of the density estimate is high. Whatever method is used, when effects are detected that exceed concern levels they will be put into perspective in the context of the entire study as well as other available information to determine if or what additional data are needed. A “no-pass” result does not necessarily mean that definitive field testing is automatically required.

(iv) For example, a test may be run in an area where a species is abundant, yet on a specific study site their numbers may be sufficiently small that a single death exceeds the level of concern on that site. Statistically, such a finding would indicate that the study did not pass according to the binomial approach, and this would be the preliminary interpretation. However, if the other sites had an adequate number of this species as well as other species expected to be at risk and no other signs of impacts are observed, the implications of the mortality would seem minor. On the other hand, if diversity of species were extremely limited, it would have greater significance. In other situations where the one dead bird is of a

species with small numbers on most sites, but density and diversity of other species is representative of nontargets expected to be at risk, another screening study that looks at the species in which the effect was detected may be appropriate. Conversely, a screening study showing that there is appreciable mortality on most study sites may be sufficient for the Agency to consider regulatory action.

(v) In summary, the interpretation of results will go beyond the statistical evaluation since the Agency must consider all the factors and circumstances peculiar to each test and site. The biological interpretation of results is, and probably always will be, a matter of scientific judgment based upon the best available data. In general, the judgmental aspects of biological interpretation are more important for definitive studies than for screening studies. Nevertheless, biological considerations often will be relevant to screening studies. Study conclusions must integrate that which is biologically significant with that which is statistically significant.

(vi) Another consideration in the interpretation of results of a field study is the attribution of effects to the pesticide being studied. A well-designed study will include appropriate techniques to determine if an effect is caused by a pesticide. In the absence of such techniques, the Agency has no choice but to consider that any effects were as a result of the pesticide use. As an example, measurement of ChE levels can provide information, since it is generally accepted that inhibition of 20 percent indicates exposure and inhibition of 50 percent or more indicates, in birds, that mortality is due to an inhibitor (see paragraph (h)(4) of this guideline). If the test chemical is the only ChE inhibitor used in the vicinity of the study site, it can be reasonably assumed that a mortality associated with 60 percent ChE inhibition is due to the test chemical. However, if other ChE inhibitors are used near the site, additional information, such as residue measurements, may be necessary to attribute death to the specific ChE inhibitor being tested.

(e) **Definitive study**—(1) **Objective and scope.** (i) The definitive study is a relatively detailed investigation designed to quantify the magnitude of impacts identified in a screening study or from other information. In contrast to the screening study, which monitors mainly the proportion of the local population that is expected to be exposed, the definitive field study examines a sample of the entire local population in the treated area. Although a definitive study may be performed when laboratory studies indicate a high potential for field mortality, it is more likely to be requested when there is evidence that actual field mortality has occurred, as in a screening study, or where reproductive effects are being investigated. The objectives of the definitive study are:

(A) To quantify the magnitude of acute mortality caused by the application.

(B) To determine the existence and extent of reproductive impairment in nontarget species from the application.

(C) To determine the extent to which survival is influenced.

(ii) Due to the intense effort and time required to estimate these parameters, the definitive study should be limited to one or a few species believed to be at the highest risk. If it can be shown that minimal (as defined at the onset of the study) or no changes occur in study parameters to high risk species, there is likely to be minimal potential for adversely affecting other presumably low risk species from use of the pesticide in question.

(iii) The definitive study, in addition to estimating the magnitude of effects of acute toxicants, also can be applied to estimating the magnitude of chronic or reproductive effects. Although the discussion has emphasized chemicals that are acutely toxic, with few exceptions it is applicable to chemicals that cause chronic effects.

(iv) In general, the definitive study will provide limited insight into whether or not effects are within the limits of compensation for the species of concern. However, using the data collected in these studies coupled with available information on the species of concern and basic theories of population dynamics, the meaning of the observed effects on the species can be evaluated.

(2) **Sampling and experimental design.** (i) The principles of statistical design of studies are well documented and it is beyond the scope of this guideline to cover the fundamentals of this topic. However, there are a few points on this topic that warrant discussion relative to the definitive study.

(ii) In the design of field studies, it is necessary to consider carefully what constitutes a sampling unit. Eberhardt, under paragraph (h)(9) of this guideline, points out that special problems are faced in designing experiments on wild animal populations. Study sites must be large in order to limit the influence of boundary effects, such as movements into and out of the area. Large study sites can be very expensive both in terms of actually applying the experimental treatment and in the assessment of results. Eberhardt also states that numerous observations, even a full year of data, on a single study site may result in very sound values for that site, but do not provide a basis for inferences to other sites. Hurlbert (see paragraph (h)(20) of this guideline) has discussed the problems associated with field studies where there was no replication or replicates were not statistically independent, which he terms pseudoreplication. Of the field studies he evaluated, 48 percent of those applying inferential statistics had pseudoreplication.

(iii) According to Eberhardt, under paragraph (h)(9) of this guideline, lack of replication seems to be based on the mistaken assumption that variances based on subsampling of sites (intrasite variability) are suitable bases for comparing treatment effects (intersite variability). This, he believes, is not a valid basis for a statistical test, because it is the variance of sites that are treated alike that is relevant to a test of treatment differences. Although subsampling of sites may be necessary to collect the data, it is the difference between sites that is important for analysis.

(iv) An important point to consider in designing a definitive study is to be sure that the study will detect a substantial impact when, in fact, it occurs. In statistical terms this concept is referred to as the power of the test. Experience with classical experimental designs with random assignment of experimental treatment and controls, has shown that the probability of a Type II error is generally high (unless very large numbers of replicates are available). Eberhardt, under paragraph (h)(9) of this guideline, indicates that, all too often in field studies on impacts to wildlife, either by default or lack of understanding, there is only a 50 percent chance of detecting an effect, which he likens to settling the issue by flipping a coin and doing no field study whatsoever. Since a definitive study is carried out under the assumption that effects will occur, the Agency believes minimizing Type II errors is extremely important.

(v) As suggested, the more generally used experimental designs require inordinately large sample sizes to obtain small Type II errors. For example, based on a coefficient of variation of 50 percent (a relatively homogeneous sample for the kinds of data collected in field studies; Eberhardt (see paragraph (h)(8) of this guideline), a 20 percent minimum detectable difference between means, a Type II error of 0.2 and a Type I error of 0.05, the number of replications required can be estimated as:

$$n = (Z_{1-\alpha} + Z_{1-\beta})^2 (CV)^2 [1 + (1 - \delta)^2] / \delta^2$$

where

n = number of replications

$Z_{1-\alpha}$ and $Z_{1-\beta}$ are critical values of the unit normal distribution

CV = coefficient of variation

δ = detectable mean difference expressed as the proportion of the control group mean, i.e., $\delta = (\mu_1 - \mu_2) / \mu_1$

For the above example:

$$n = (1.65 + 0.84)^2 (0.5)^2 [1 + (1 - 2)^2] / (0.2)^2$$

$$n = 63.6 \text{ replicates}$$

Thus, for the above parameter, 64 replicates for both control and treatment groups, or 128 total study plots, are required to detect a 20 percent difference between treatments and controls with an 80 percent chance of being sure to detect a real difference at a 0.05 level of significance.

(vi) With more sophisticated designs, the number of replicates can be reduced under some circumstances and still meet the Agency's aim to limit the probability of a Type II error to 0.2 with a detectable difference of 20 to 25 percent. For example, a paired plot design can be used, substantially reducing the number of replicates required. Pairing serves to reduce the effective coefficient of variation by reducing the variation attributable to experimental error. The lower coefficient of variation reduces the number of replicates. A quantitative difference or, preferably, a ratio of treated to the total of treated and control responses, can be analyzed statistically to test for a treatment effect on the measured response variables (see paragraph (h)(31) of this guideline). The logic of using paired plots is that, while no two areas are ever exactly alike, two areas that are not widely separated in space are ordinarily subjected to much the same climatic factors, have populations with about the same genetic make-up, and generally the two populations can be expected to follow much the same trend over time, apart from a pesticide effect (see paragraph (h)(8) of this guideline)). If all plots are approximately equal in area and habitat and population densities between pairs are similar, it is postulated that when no pesticide impacts occur, the mean ratio of treatment to treatment plus control will equal one-half. Then a t-test or an exact randomization test (see paragraph (h)(11) of this guideline) may be applied to test whether the average number of survivors on the treated plots is equal to the average number of survivors on controls.

(vii) The number of pairs required can be estimated using the following formula:

$$n = (Z_{1-\alpha} + Z_{1-\beta})^2 \times [4 q_1 / (p_1)^2 (1 + q_i) c]$$

where

n = number of paired plots

$Z_{1-\alpha}$ and $Z_{1-\beta}$ are critical Z scores

q_1 = survival ratio

p_1 = mortality ratio

c = mean number of survivors on control plots

(viii) Therefore, at an 80 percent assurance of detecting a treatment-induced impact of 20 percent or greater at a 0.05 level of significance if c = 28,

$$n = (1.65 + 0.84)^2 \times 4(0.8) / [(0.2)^2(1 + 0.8)(28)]$$

$$n = 9.84$$

Thus, 10 pairs of plots (20 total) with a mean of 28 individuals per plot would be needed. Increasing the mean number of individuals per plot (c), causes a reduction in n.

(ix) In some field situations, pairing may not be feasible. In these situations, other designs would be more appropriate or a less rigorous design may have to be used. However, in planning field studies, the power of the study design must be considered to determine the limitations of the study. Studies with adequate replication are highly preferred to support registration—the use of less replication will not necessarily render the study inadequate. However, it is wrong to use a study with low power to imply no biological damage, when the study is not capable of detecting the damage if it occurred. In cases where large numbers of replicates are impractical, subjective and biological knowledge should be used in a decision process to decide if there was a treatment effect. In most instances, it is highly advisable to involve statisticians or biometricians who are familiar with this kind of field study in the planning and analysis phase of the field work to avoid costly technical errors.

(3) **Study area and site selection.** Selection of geographical areas and study sites within the areas for the definitive test generally requires the same considerations as for a screening study. For the definitive study, however, the selected areas and study sites must have adequate populations of the species of concern. Obviously, the crop of concern must be grown on a representative portion of the area. In addition, consideration needs to be given to whether the target pest species will be present. If it is not, it is necessary to consider what influence its absence may have on potential results. For example, if the pest is a major food source for nontarget species, its absence could significantly influence results. Finally, the potential variation in populations of concern over the geographical areas selected should be considered. It may be difficult to find sites that are sufficiently similar to provide paired plots, which limits the coefficient of variation so that the desired sensitivity can be achieved.

(4) **Number and size of sites.** As suggested under paragraph (e)(2) of this guideline, the number of sites will depend upon the species density on sites and the sensitivity required. Ideally, sample size should be large enough so there will be an 80 percent probability of being sure to detect a 20 percent difference when it exists. The size of the study site must be large enough to provide adequate samples. The size depends on the survey methods used, sensitivity required, and the density and range of the species of concern. For a paired-plot design, the number of sites required is a function of the average density of the species. In general, the breeding density of the species of concern can be used to provide a rough

estimate of the size of area needed to provide adequate samples. However, preliminary sampling most likely will be required to verify the estimates.

(5) **Methods.** (i) Essentially, the methods used in a definitive study are a means to quantitate reproductive and mortality rates of animals on treatment and control areas. There are many texts and monographs available on methods of sampling to estimate these parameters. Anyone not familiar with the theory and principles of the various techniques should review these references in depth. The objective of this portion of the guideline is to provide a general guide to the various methods that could be used in a definitive field study. In addition, these methods can be applicable to some screening studies.

(ii) The methods to be used in an individual field study will depend on the nature of the identified concerns. Some methods are useful for investigating several types of concerns, and most types of concerns can be studied by several methods. When the concern becomes more specific (e.g., secondary hazards to raptors) or the use pattern and/or habitat type is limited, the range of applicable methods tends to become more narrow.

(iii) Methods described below are divided into three categories: Methods for assessing mortality and survival of adults and independent juveniles, methods for assessing reproduction and survival of dependent juveniles, and ancillary methods. The intent of this guideline is to present methods that are likely to be useful in many situations, rather than an exhaustive list of all available methods. The Agency encourages the use of other methods when they are scientifically valid and have a high probability of detecting an effect.

(iv) While it is absolutely essential to have a detailed plan that describes the selected actions (with contingencies) for achieving the study objectives, investigators must remain flexible because unanticipated problems always come up in long term studies. Even with highly experienced and resourceful field biologists, the most carefully planned studies can be compromised due to the unpredictability of wild animals and natural events. When a natural disaster occurs early in the study, it may be wise to reinitiate the study. If the event occurs after substantial data already have been collected (e.g., early in the second year of a multiyear study), it may be more appropriate to extend the study an additional year or more to help provide for the additional needs. If the study is to be terminated, the report should describe thoroughly the nature of the events and the consequences if they affect the study results.

(6) **Mortality and survival.** (i) It is very important to understand the autecology of the species being studied in order to select the most appropriate methods for investigating them. In addition, the choice of particular methods must consider the applicability of the method based on the pesticide use pattern and study site characteristics.

(A) **Mark-Recapture.** (1) There are several mark-recapture methods available, each based on the same basic premise. A sample of animals is captured, marked, released, and another sample is collected where some of the animals are captured again. The characteristics of this identifiable sample are used to estimate population parameters. Mark-recapture studies can provide information on:

- (i) Size of the population.
- (ii) Age-specific fecundity rates.
- (iii) Age-specific mortality rates.
- (iv) Combined rates of birth and immigration.
- (v) Combined rates of death and emigration.

(2) Seber (see paragraph (h)(32) of this guideline) reviewed the various mark-recapture methods and subsequent statistical analyses. Less detailed but still very useful reviews are provided in paragraphs (h)(4) and (h)(16) of this guideline. Nichols and Pollock provide a valuable comparison of methods under paragraph (h)(26) of this guideline. The following Table 2. provides a brief summary of some of the various mark-recapture methods discussed in these references.

Table 2.—Mark-Recapture Techniques

Method	Applications/Requirements/Assumptions
Peterson Method (Lincoln Index)	Estimation of population size. Usually only two sampling periods. Closed population.
Schumacher's Method	Estimation of population size. More than two sampling periods. Marking continues throughout sampling. Closed population.
Bailey's Triple Catch	Estimate of birth rate and death rate in addition to population size. Requires data from two marking occasions and two recapturing occasions. Open population.
Jolly-Seber Method	Estimates mortality and recruitment in addition to population size. Requires more than two sampling periods and that each animal's history of recapture be known. Open population.

(3) When considering the use of one of these mark-recapture models, one must carefully evaluate the applicability of the method to the circumstances under consideration. While in theory mark-recapture techniques should be an excellent method for evaluating effects of pesticides on wildlife populations, some mark-recapture analyses are not particularly robust; small deviations from their implicit assumptions can produce large errors in the results (see paragraph (h)(4) of this guideline). However, some of the more recent and sophisticated analytical methods are robust and can deal with deviations from assumptions in closed populations (see paragraph (h)(28) of this guideline).

(4) Mark-recapture methods are particularly useful for small mammals because these animals are seldom amenable to the visual and auditory observations necessary for using transect, territory mapping, or similar methods. However, mark-recapture also may be useful for birds provided a sufficient number of birds can be captured and marked. In some situations, birds may be “recaptured” with use of binoculars via visual observations of marked individuals. Animals must remain marked for the duration of the study. Typically, mammals are toe-clipped or ear-marked and birds are banded. Marking should not make the animals more susceptible to the effects of the pesticide (e.g., anticoagulants with toe clipping). Dyes may be useful unless they are lost by wear or molting.

(B) Territory mapping method. (1) A common spatial census method is territory mapping, wherein the territories of individuals are mapped before and after treatment, on both treated and untreated plots. The method is usually applicable when birds are defending territories. It involves a series of census visits to the study sites during which birds located by sight or song are recorded on a map. The information from all the visits is plotted for each species. Birds exhibiting territorial behavior appear on the map as clusters of individual contacts. The clusters are used to estimate both the size and number of territories. The pre- and posttreatment censuses for treated sites are compared with the pre- and posttreatment censuses for control sites to determine changes in populations of territorial individuals that may be attributed to the pesticide (see paragraph (h)(12) of this guideline). Further details of this method are given by the International Bird Census Committee under paragraph (h)(21) of this guideline, and its application to evaluating impact caused by pesticides is reviewed by Edwards et al. (see paragraph (h)(12) of this guideline).

(2) Problems with this method can occur. Under some circumstances, replacement from outside the area can be so rapid that territories are re-filled before the census is completed. There usually is a floating population of silent, nonterritorial birds who may quickly reoccupy empty territories (see paragraph (h)(35) of this guideline). The effects of replacement can be overcome for some species by capturing and marking the territorial individuals prior to treatment, so they can be distinguished from the floaters. Also, replacement may not be a problem when the study areas are in the center of a relatively large treated area.

(C) Radio telemetry. Radio telemetry can be an extremely useful technique to provide information on the effects of a pesticide application on nontarget species. As discussed for screening studies, radio telemetry can be used to monitor for mortality as well as to provide useful information on behavioral modification caused by the pesticide application. The points discussed previously (for screening studies) generally are applicable to definitive studies. However, for the definitive study, the number of radio-tagged animals needed depends upon the variation between sites and the sensitivity required. For example, with behavioral observation, intra-

and intersite variation will influence the number of radio-tagged animals required. In some instances, it might not be practical to radio-tag the number of animals required to provide a rigorously designed study. Under these conditions, the limitations should be specified, and the maximum number of animals that can be practically radio-tagged and monitored should be used.

(D) **Other methods for mortality and survival.** Other techniques for assessing density and diversity are discussed for screening studies; most of these, especially line-transect methods, are useful for definitive studies. Some methods, such as catch per unit effort or counts of animal signs, do not provide actual measures of density but may still be used to compare effects on treated and untreated plots.

(7) **Reproduction and survival of dependent young.** Some of the techniques for assessing mortality and adult survival are also useful for assessing reproduction and survival of young. Some, but not all, mark-recapture methods can provide information on fecundity. Radio-tagging nestlings or suckling young of moderate and large size animals may be used to assess survival of dependent young. Radio telemetry and territory mapping are useful for locating dens or nests for further study. The following methods are more specific for assessing reproductive parameters.

(i) **Nest monitoring.** (A) Nest monitoring is useful for evaluating the effect of pesticides on breeding birds. The typical procedure is to search the study site to find active nests and subsequently to check those nests to determine their fate. Information collected on each nest should include number of eggs laid, number hatched, number of young fledged, and if and when the nest was abandoned or destroyed, both before and after pesticide application. While all definitive studies should consider this technique, it also may be useful in screening studies.

(B) This technique is relatively straightforward. However, it may not be practical if nests are scarce or otherwise hard to find. Because the breeding success of birds can be highly variable and can be quite low, it is sometimes difficult to obtain sufficient data on the success of the same species in enough sites to yield satisfactory results for statistical comparison with controls (see paragraph (h)(18) of this guideline). In some cases, artificial nest structures can be constructed to increase nest densities. In a few situations where sufficient numbers are available, the technique may be applicable to mammal den monitoring.

(ii) **Behavioral observations.** Behavioral observations associated with reproduction can be quite useful, especially for birds. Techniques are simple, but labor intensive. When used, such observations most likely would be combined with nest monitoring since both techniques require locating reproductive sites. Typically, the frequency and duration of behaviors will be compared for treated and untreated plots. Incubation, parental

care (especially feeding for altricial birds), and following behavior (for precocial animals) are behaviors that are particularly amenable to such study. Courtship, mating, and nest building are other behaviors that could be studied in some situations, but locating sufficient numbers of animals displaying these behaviors to permit quantitative analysis is difficult.

(iii) **Age structure of populations.** (A) Comparisons of young/adult ratios of selected species between treated and untreated plots may indicate reproductive effects. The timing of the application and of breeding of selected species is critical. For assessing reproductive impairment or survival of dependent young, per se, the duration of this technique should be limited to single breeding/rearing periods, which may be repeatedly assessed. However, longer study periods that may even include several years can be used to assess the combination of reproductive success and age-specific mortality, even if the two cannot be separated.

(B) Obviously, use of this method requires that the age of individual animals be determined. In some cases, it may be necessary only to distinguish among adults, subadults, and juveniles. In mammals, this may usually be accomplished by examining pelage, development of testes or mammarys, or tooth eruption or wear characteristics. In birds, plumage or characteristics of particular (species-dependent) feathers may be used. For carcasses or sacrificed animals, observation of the ossification of bones or development of reproductive organs are useful. In other cases, particularly where comparisons are made among populations in different years, it may be appropriate to distinguish age classes of adults. In mammals, tooth eruption, wear, or enamel layers, or eye lens weights are useful. It is more difficult to separate age classes of many adult birds, although overall plumage or feather characteristics can provide some indication. In some slow-maturing birds (e.g., gulls), plumage may be used to distinguish year classes of sub-adults. Additional details on aging birds and mammals are presented by Larson and Taber (see paragraph (h)(22) of this guideline).

(8) **Ancillary methods.** (i) At least some ancillary methods are essential in every field study. As used here, ancillary methods are generally of two types. Certain of these methods are important for determining the nature or existence of effects or for establishing causal relationships. Others of these methods do not address effects directly, but they provide important information for interpreting the results of the study.

(ii) Many of the methods for determining effects have been discussed for screening studies. Enzyme analysis, such as for ChE inhibition, and observations of signs of toxicity can show that animals were exposed to or killed by a toxicant of a particular type. Where it is possible that animals may be exposed to other pesticides of the same type (e.g., feeding in a nearby area treated with other pesticides), residue analysis in nontarget animals may be necessary to determine which specific pesticide caused

the signs or alterations in enzymes. Even though carcass searches, per se, are not recommended for definitive studies, it is still essential to recover and analyze any carcasses found accidentally or obtained through radio-tracking. Residue and/or enzyme analysis of live animals collected will frequently be important.

(iii) Among the other ancillary methods, analysis of environmental residues is crucial and will probably be necessary in nearly every definitive field study. As discussed for screening studies, the most important environmental residues are those that occur on or in wildlife food sources, which may include insects, plant parts, or even other vertebrates, depending upon the species that are the primary focus of the investigation. The investigator should review the literature on food habits of the species being studied; often it will be appropriate to assess food habits on the specific study sites, particularly where the literature is not adequate to define food habits in the agricultural ecosystem under study. Such an assessment should include the availability of food sources and the number of mobile animals that spend only part of the time in and adjacent to treated sites. The habitat should be thoroughly described to include both the morphology and species that are relevant to wildlife. Frequently, it will be important to locate and describe roosting, denning, or nesting sites for mobile wildlife that use treated sites part of the time.

(9) **Interpretation of results.** (i) While each field study is unique, some elements may be common among many field studies. When a definitive field study is required, the requirement is based on one or more specific concerns that pertain to a specific chemical and one or several use patterns. Because of the substantial diversity in the types of problems to be assessed and the variety of available investigative methods, the key to understanding and interpreting a field study lies in the development of a sound protocol. All protocols will contain a description of the study sites, or the characteristics to be used in selecting sites within a given area, and the methods to be used in conducting the study. However, a well designed protocol will go beyond this descriptive approach in three ways.

(A) First, the well-designed protocol will contain a restatement of the concerns to be addressed to ensure that there is an adequate understanding of the Agency's position. The investigator should review the literature and other available information that may bear upon the problem. It is possible that the literature may contain a valid answer to the questions raised by the Agency. Far more likely, the literature may orient the investigator to address the concerns in a particular way. An example is provided by Hegdal and Blaskiewicz (under paragraph (h)(17) of this guideline), who conducted a study to address the Agency's concerns for secondary toxicity to barn owls (specifically) from the use of an anticoagulant bait proposed for use on commensal rodents in and around agricultural buildings. A review of the literature by these investigators indicated to them that labora-

tory studies suggested a legitimate potential for secondary poisoning to exposed raptors, but that the food habits of barn owls consist primarily of microtine rodents in most areas, suggesting a low potential for actual exposure. Consequently, they designed their study to focus on barn owl food habits and movements, and included an additive to the bait formulation that would permit an identification of whether or not the barn owls ate rodents that had fed on the bait. The study adequately demonstrated that actual exposure of barn owls was quite limited, and the proposed registration for this use was subsequently approved. By using the available literature on both the chemical and the particular species of concern, the investigators were able to narrow the study while still providing sufficient information for evaluation. However, it should be noted that this study was not adequate for evaluating the potential for secondary toxicity in the field to other predators that may have different food habits, or for other use patterns that may result in exposure to different predators or scavengers.

(B) Second, the well designed protocol will provide the reasons why particular methods are being used, including, at least qualitatively, the meaning that different results might have. For example, a protocol may include collection of residues in nontarget animals, but it also should include a statement of purpose and meaning for such collection. Residues may be used to indicate potential exposure to nontarget organisms through analysis of their food, exposure in nontarget animals as a result of eating contaminated food, or that a particular pesticide was likely to be the cause of any observed effects. Interpretation of data is facilitated substantially by a statement of what results were intended by using a particular technique. In the previously cited example (see paragraph (h)(16) of this guideline), it was clearly stated that collection of owl pellets was used to assess general food habits and that use of a fluorescing dye in the bait was for the purpose of ascertaining whether or not the owls fed on commensal rodents that specifically had fed on the bait. The interpretation of the data collected, once the purpose was stated, naturally led to the conclusion of no-significant-exposure to the barn owls.

(C) Third, the well designed protocol will contain an experimental design that will indicate how the results can be assessed quantitatively. The experimental design has been discussed in previous sections of this guideline, but there are two facets that relate closely to the interpretation of results: The difference that can be detected between treated and untreated plots and the power (ability) of the design to detect this difference. An experimental design with number of replicates based on an estimated coefficient of variation that closely approximates reality will allow the study to detect a stated concern level some prescribed number of times during the study time. The actual difference between treated and control units is measured during the field study, but the design can form an initial basis for interpretation when combined with the available information on

the species of concern. As a result, the well-designed protocol should include a section on interpretation.

(ii) Study methods for investigating acute mortality are more straightforward than for other kinds of effects. Nevertheless, there are sufficient differences in the use of the data to preclude a constant interpretation. The study may focus directly on the species of concern and may involve little or no extrapolation, depending on such factors as the type and the extent of use, the available toxicity data base, and home range of the species. Extrapolation to other populations, regions, or uses might be necessary. If the species of concern cannot be studied directly, it may be necessary to extrapolate between species, involving interspecies differences both in toxicological sensitivity and in ecological and population parameters.

(iii) The same kinds of considerations apply to reproductive impairment and chronic toxicity, even though different, and often more laborious and costly, investigative methods are involved. Where reproductive success is impaired, information on species-specific variation in reproductive ecology is necessary to understand how a particular degree of impairment may relate to effects among various species. Such reproductive considerations can include whether an avian species is a determinate or indeterminate layer, the number of nestings per season for different geographic areas in the use pattern, the length of the refractory period, as well as the specific effect which can range from destruction of reproductive organs to behavioral deficits such as nest abandonment. Considerations of reproductive ecology among different species of mammals include delayed fertilization or implantation, resorption of embryos or parental infanticide due to stress, number of young per breeding cycle, etc. All of these factors, and many others, are relevant to determining for different species the extent of effects that could result in population reductions or lack of ability to recover.

(iv) An analysis of whether or not a particular level of effect is going to affect wildlife populations is species-specific. For any species (or subspecies), the changes in population can be described very simplistically by the equation: rate of population increase (r) = birth rate – death rate, where values of r can be positive (population growth) or negative (population reduction). Immigration and emigration are also important when the concern is for specific populations of a species. These characteristics differ among species, and data will not always be available. The application of sound scientific judgment to the best available information will be the basis for interpreting the results of a study. It may be necessary to compare the results of the field study to laboratory data, especially where laboratory data are available on a variety of species and/or effects and the field study has focused on species other than those of direct concern. The use of extrapolation techniques will be necessary where endangered species are of concern or where other species cannot be studied directly.

(v) The Agency would like to be able to obtain a standardized result from a field study so that the result could be applied in a very consistent manner. As discussed in previous sections of this guideline, the different effects and species of concern will vary and will require the development of specific protocols to address these factors. Although most of the various techniques have some degree of standardization, the field study may combine the individual techniques in a wide variety of ways to address specific concerns. A standardized result might be attainable for the individual techniques, although that result would still have to be applied differently for various species, depending on their biology and ecological characteristics. However, determining a result for the whole field study that would unequivocally lead to a statement of the degree of risk, while obviously desirable, is not currently practical.

(f) **Suggested components of a field study protocol.** The following protocol was adapted from The Wildlife Managements Techniques Manual (under paragraph (h)(29) of this guideline), and is recommended for studies submitted to the Agency for review.

(1) Title.

(2) Problem definition. The following information should be provided:

(i) A review and summary of the available information on the pesticide in relation to nontarget hazard, including use information.

(ii) A precise statement of the goals and purpose of the study.

(iii) A brief statement of the problem and the context in which it exists, specifying the limits of the proposed work.

(iv) Precise statements of the major hypotheses to be tested.

(3) Methods and materials. This section should include the following:

(i) A brief discussion of various methods and procedures that have been or could be used to evaluate the problem. This discussion should identify the strengths and weaknesses of each method or procedure discussed.

(ii) Description of the procedure to be followed:

(A) Identify the study areas selected and their general suitability for achieving the objectives of the study or what criteria will be used to select study areas.

(B) Identify the species present or expected to be present on the study areas, discussing characteristics pertinent to the problem being evaluated.

(C) State the research procedures, designs and sampling plans to be used.

(1) Specify the kind and amount of data needed and to be sought.

(2) Describe in detail how all the data are to be obtained, including details of application, instrumentation, equipment, sampling procedures, and any other information.

(D) Describe how the data are to be treated, including specifying what statistics are to be calculated, what models will be used, what tests of data will be used, etc.

(E) Describe in detail the methods to be used to check the sensitivity and accuracy of the procedures used.

(F) Describe quality assurance procedures for application, instrumentation, equipment, and records.

(G) Briefly describe the resources (people, facilities, etc.) to be applied to the study.

(g) **Carcass searches**—(1) **Design.** (i) In designing carcass searches, the following factors need to be known or determined:

(A) Density of the species that are likely to be exposed. For example, granular products are most likely to result in exposure to ground-feeding animals; therefore, birds such as warblers or swallows should not be included in density counts for such products.

(B) Probability of finding dead animals if any are killed. This is dependent on the probability of a carcass remaining on the study site (i.e., not being removed by scavengers) and the probability of detecting a carcass if it remains on the study site (search efficiency).

(C) Size of the search area.

(D) Number of carcasses found.

(ii) These factors can be combined in the following formula:

$$n = d r e a p$$

where n = number of carcasses found, d = density in animals per acre, r = proportion of carcasses remaining (nonremoval), e = search efficiency, a = acres searched, and p = proportion of population killed.

(iii) Carcass searches should be used only when there is a reasonable potential to detect mortality. If such mortality does occur, the carcass search should be able to detect it and therefore, carcasses should be found. It is recommended that carcass searches be designed so that at least two carcasses ($n = 2$) will be found if there is appreciable mortality. In general,

preliminary sampling would be required to determine these factors. However, information from other field studies can be used in the planning stages to determine if carcass searching would be appropriate for use under anticipated conditions and to assist in developing the study design.

(iv) The sensitivity of the carcass search approach is equivalent to the percent detectable kill of the population. To determine the sensitivity, the formula is adjusted:

$$p = n / d r e a$$

Since p is a proportion:

$$\text{percent detectable kill} = p \times 100 = n / d r e a \times 100$$

(v) If any of the values of d , r , e , or a are zero, the equation cannot be solved and the carcass search is not applicable (i.e., no density of birds, no acres searched, no carcasses remaining, no remaining carcasses found). However, other combinations of d , r , e and a , such as low density and small acreage or low efficiency and high scavenger removal, can result in a small denominator meaning that mortality can be detected only when a high percentage of the population is killed. For example, in 5 acre fields with only two birds per acre and r and e estimated at a moderate 0.5, only an 80 percent or greater kill could be detected. In such situations, it is necessary to increase one of the parameters to achieve a stated level of detectability or else to use methods other than carcass searching. The same equation can be used to estimate the minimum search area to detect a given mortality level (p) by solving for a .

(2) **Search procedure.** In general, depending on the sensitivity of the search method relative to the habitat involved, corridors or plots should be selected. These areas should be searched systematically by walking predetermined routes until the area has all been covered. Due to the concentration required to find dead animals, other activities that could distract the attention of the searchers should be avoided during carcass searching. In homogeneous situations, investigators should randomly select search areas. However, in most studies it is advisable to stratify the sampling, concentrating efforts in areas frequented by wildlife species such as woods edges, ditch banks, field borders, fencerows, and other habitats where wildlife concentrate.

(3) **Duration.** Searches should begin on the day of application and continue on a daily basis for as long as mortalities or other evidence of intoxication occur. In general, a week or two following application should be adequate. However, the length of time searches are continued should be related to how long lethal concentrations are expected to be present. Normally, the same areas should be searched each day.

(4) Estimating efficiency of carcass search. (i) Efficiency trials should be conducted periodically (minimum 3 times per study site) during the study to determine the proportion of carcasses that are detected. Just prior to the initiation of a scheduled search, carcasses of animals representative of species found in the area should be variously placed within the search area. If the study site includes edge habitat, carcasses should be placed in the edges as well as in the fields. In general, carcasses should be placed where animals would be most likely to die, depending on the nature of the chemical. Searchers should not be aware that simulated mortalities have been placed; however, they should be aware that these trials will occur during any scheduled search.

(ii) The number of carcasses placed should be approximately equal to 20 percent of the estimated density of species on the search area. All placed carcasses should be marked to distinguish them from actual kills. The location of placed carcasses should be mapped so those not found can be easily recovered following completion of that day's search activities, since unrecovered carcasses could bias study results. For example, if a scavenger were to carry off a simulated kill and consume it at another location on the study site, the remains could be erroneously classified as pesticide-related if found. One way to avoid this problem would be to dip carcasses in a nontoxic substance that fluoresces under ultraviolet light so that the remains could be identified.

(5) Estimating carcass removal rate. (i) Carcass removal should be monitored to determine local variability in scavenger activity. The density of both carcasses and scavengers can influence the rate of removal. Under some conditions, large numbers of carcasses may attract scavengers. In other situations a large number of kills may dilute removal rate due to limited number of scavengers. Where it can be adequately documented that removal of carcasses occurs almost exclusively either at night or during the day, the timing of carcass searches may be adjusted to minimize the effects of removal.

(ii) Carcasses planted in monitoring trials should simulate mortalities actually occurring from the pesticide. In most cases, small to moderate sized species such as starlings or blackbirds, or laboratory bobwhite or Japanese quail chicks may be used. Carcasses should be variously placed within the general study areas and monitored daily for at least 5 days or until 90 percent have been removed. The number used should approximate densities resulting from effects of the pesticide under study; however, in most instances, this will not be known. Therefore, a density of approximately 20 percent of the population of nontarget species on the area is recommended.

(iii) Timing of carcass removal trials should be such that they do not affect scavenger removal of pesticide-killed birds or the feather-spots of the removed carcass could be erroneously classified as a pesticide kill.

Location of placed birds should be recorded on maps and may be marked in the field with small stakes or by other inconspicuous means, preferably at a fixed distance and direction from the carcass.

(h) **References.** The following references should be consulted for additional background material on this test guideline.

(1) Bunyan, P.J. et al. Organophosphorus poisoning, some properties of avian esterases. *Journal of Agricultural and Food Chemistry* 16: 326–331 (1968).

(2) Bunyan, P.J. et al. Organophosphorus poisoning, diagnosis of poisoning in pheasants owing to a number of common pesticides. *Journal of Agricultural and Food Chemistry* 16: 332–339 (1968).

(3) Burnham, K.P. et al. Estimation of density from line transect sampling of biological populations. *Wildlife Monographs*, No. 72 (1980).

(4) Caughley, G. *Analysis of Vertebrate Populations*. Wiley, NY (1977).

(5) Cochran, W.W., Wildlife telemetry. pp. 509–520, in *Wildlife Management Techniques Manual*. S.D. Schemnitz, Ed. The Wildlife Society, Washington, DC (1980).

(6) Corbett, J.R. *The Biochemical Mode of Action of Pesticides*. Academic, NY (1974).

(7) Davis, D.E. and R.L. Winstead. 1980. Estimating the numbers of wildlife populations. pp. 221–246, in *Wildlife Management Techniques Manual*. S.D. Schemnitz, Ed., The Wildlife Society, Washington, DC (1980).

(8) Eberhardt, L.L. Quantitative ecology and impact assessment. *Journal of Wildlife Management* 4: 27–70 (1976).

(9) Eberhardt, L.L. Appraising variability in populations studies. *Journal of Wildlife Management* 42: 207–237 (1978).

(10) Eberhardt, L.L. Assessing the dynamics of wild populations. *Journal of Wildlife Management* 49: 997–1012 (1985).

(11) Edgington, E. *Randomization Test*. Dekker, NY (1980).

(12) Edwards, P.J. et al. The use of a bird territory mapping method for detecting mortality following pesticide application. *Agro-Ecosystems* 5: 271–282 (1979).

(13) Ellman, G.L. et al. A new and rapid calorimetric determination of acetylcholinesterase activity. *Biochemical Pharmacology* 7:88–95 (1961).

- (14) Environmental Protection Agency. Pesticide Assessment Guidelines, Subdivision E – Hazard Evaluation: Wildlife and Aquatic Organisms. Office of Pesticides and Toxic Substances, Washington, DC (1982).
- (15) Hanson, W.R. Estimating the density of an animal population. *Journal of Research Lepidoptera* 6: 203–247 (1967).
- (16) Hegdal, P.L. and R.W. Blaskiewicz. Evaluation of the potential hazards to barn owls of Talon (brodifacoum bait) used to control rats and house mice. *Journal of Environmental and Toxicological Chemistry* 3: 167–179 (1984).
- (17) Heinz, G.H. et al. Environmental contaminant studies by the Patuxent Wildlife Research Center, pp. 8–35. in *Avian and Mammalian Wildlife Toxicology*. ASTM STP 693, E.E. Kenaga (Ed), American Society for Testing and Materials, Philadelphia, PA (1979).
- (18) Hill, E.F. and W.J. Fleming. Anticholinesterase poisoning of birds: field monitoring and diagnosis of acute poisoning. *Journal of Environmental and Toxicological Chemistry* 1: 27–38 (1982).
- (19) Hurlbert, S.H. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54:187–211 (1984).
- (20) Giles, R.H. Jr. *Wildlife Management*. Freeman, San Francisco, CA. (1978).
- (21) International Bird Census Committee. Recommendations for an international standard for a mapping method in bird census work. *Bulletin of the Ecological Research Committee* 9:49–52 (1970).
- (22) Larson, J. S. and R. D. Taber. Criteria of sex and age. pp. 143–202, in *Wildlife Management Techniques Manual*. S. D. Schemnitz; (Ed.), The Wildlife Society, Washington, DC (1980).
- (23) Leopold, A. *Game Management*. Scribner's, NY (1933).
- (24) Ludke, J.L. et al. Cholinesterase (ChE) response and related mortality among birds fed ChE inhibitors. *Archives of Environmental Contaminant Toxicology* 3:1–21 (1975).
- (25) Moen, A.N. *Wildlife Ecology: An Analytical Approach*. Freeman, San Francisco, CA (1973).
- (26) Nichols, J.D. and K.H. Pollock. Estimation methodology in contemporary small mammal capture-recapture studies. *Journal of Mammal* 64:253–260 (1983).
- (27) O'Brien, R.D. *Insecticides: Action and Metabolism*. Academic Press, NY (1967).

- (28) Otis, D.L. et al. Statistical interference from capture data on closed animal populations. *Wildlife Monographs* 62:1–135 (1978).
- (29) Ripley, T.H. Planning wildlife management investigations and projects, pp. 1–6. in *Wildlife Management Techniques Manual*. S.D. Schemnitz (Ed.), The Wildlife Society, Washington, DC (1980).
- (30) Rosene, W. Jr. and D.W. Lay. Disappearance and visibility of quail remains. *Journal of Wildlife Management* 27:139–142 (1963).
- (31) Scientific Advisory Panel. Final Scientific Advisory Panel subpanel's report on the January 7–8, 1987 meeting concerning terrestrial field studies. Environmental Protection Agency, Washington, DC (1987).
- (32) Seber, G.A.F. *The Estimation of Animal Abundance and Related Parameters*. Macmillan, NY (1982).
- (33) Shellenberger, T.E. et al. The comparative toxicity of organophosphate pesticides in wildlife. pp. 205–210, in W.B. Diechmann, (Ed.), *Pesticide Symposium*, Halos, Miami, FL (1970).
- (34) Snedecor, G.W. and W.G. Cochran. *Statistical Methods*. Sixth Edition. The Iowa State University Press, Ames, IA (1967).
- (35) Stewart, R.E. and J.W. Alrich. Removal and repopulation of breeding birds in a spruce-fir forest community. *The Auk* 68:471–482 (1951).
- (36) Walpole, R.E. and R.H. Myers. Probability and Statistics for Engineers and Scientists. *Macmillan*, NY (1972).